

THE ASYMPTOTIC BERRY-ESSEEN CONSTANT FOR INTERVALS

TODOR DINEV AND LUTZ MATTNER

ABSTRACT. The asymptotic constant in the Berry-Esseen inequality for interval probabilities is shown to be $\sqrt{2/\pi}$.

1. INTRODUCTION

Let \mathbb{R} denote the real numbers, \mathbb{Z} the integers, and \mathbb{N} the integers ≥ 1 . Let $\text{Prob}(\mathbb{R})$ be the set of all (probability) laws on \mathbb{R} , δ_x the unit mass at $x \in \mathbb{R}$, and $N_{0,1}$ the standard normal law with distribution function Φ and density $\varphi := \Phi'$.

For $P \in \text{Prob}(\mathbb{R})$ and $s \in [1, \infty[$, let $\mu(P) := \int x P(dx)$ and $\beta_s(P) := \int |x - \mu(P)|^s P(dx)$ if $\int |x| P(dx) < \infty$, $\beta_s(P) := \infty$ otherwise, $\sigma^2(P) := \beta_2(P)$, and $\alpha(P) := \int (x - \mu(P))^3 P(dx)$ if $\beta_3(P) < \infty$. If P is a lattice law, let $h(P) := \sup \bigcup_{a \in \mathbb{R}} \{\eta \in]0, \infty[: P(a + \eta\mathbb{Z}) = 1\}$, otherwise set $h(P) := 0$. Let us put

$$\mathcal{P}_s := \{P \in \text{Prob}(\mathbb{R}) : 0 < \beta_s(P) < \infty\}.$$

In the following, we shall consider sequences $(X_k)_{k \in \mathbb{N}}$ of independent and identically distributed real-valued random variables. P will then stand for the law of X_1 and shall, for the sake of brevity, usually be omitted in the quantities μ , σ^2 , α , β_s , and h just defined. Let $P_n \in \text{Prob}(\mathbb{R})$ be given by $P_n(A) := P^{*n}(\sigma\sqrt{n}A + n\mu)$ and set $F_n(x) := P_n(]-\infty, x])$ for $x \in \mathbb{R}$. Thus P_n is the law of the standardized sum of X_1, \dots, X_n and F_n its distribution function. Now the classical Berry-Esseen theorem states the finiteness of the **Berry-Esseen constant**

$$c_{\text{BE}} := \sup_{\substack{P \in \mathcal{P}_3 \\ n \in \mathbb{N}}} \frac{\sigma^3(P)}{\beta_3(P)} \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)|.$$

It is known that $0.4097 < (\sqrt{10} + 3)/(6\sqrt{2\pi}) \leq c_{\text{BE}} < 0.4748$, with the lower bound due to Esseen [2] as an obvious consequence of his determination of the **asymptotic Berry-Esseen constant**

$$c_{\infty, \text{BE}} := \sup_{P \in \mathcal{P}_3} \frac{\sigma^3(P)}{\beta_3(P)} \lim_{n \rightarrow \infty} \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| = \frac{\sqrt{10} + 3}{6\sqrt{2\pi}},$$

and with the more recent upper bound due to Shevtsova [4].

The Kolmogorov distance used above is the supremum distance over the system of all unbounded intervals. We propose to use instead the system \mathfrak{I} of all intervals whatsoever, which is a special case of the system of all convex sets used in the multidimensional

2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 60E15.

Key words and phrases. Berry-Esseen inequality, interval probabilities, asymptotic constant.

This research was partially supported by DFG grant MA 1386/3-1.

File name: 2012-04-02-berry-esseen_for_interval_probabilities.tex.

setting. Due to the continuity properties of probability measures, it suffices to consider intervals of the form $]a, b]$, that is, we have

$$(1) \quad \sup_{I \in \mathfrak{J}} |P_n(I) - N_{0,1}(I)| = \sup_{x, y \in \mathbb{R}} |F_n(x) - F_n(y) - \Phi(x) + \Phi(y)|.$$

An obvious upper bound for the right hand side is $2c_{\text{BE}}\beta_3/(\sigma^3\sqrt{n})$. While we do not know whether here $2c_{\text{BE}}$ can be replaced by a smaller constant, we will determine below the **asymptotic Berry-Esseen constant for interval probabilities** as

$$(2) \quad c_{\infty, \text{BE}}(\mathfrak{J}) := \sup_{P \in \mathcal{P}_3} \frac{\sigma^3(P)}{\beta_3(P)} \lim_{n \rightarrow \infty} \sqrt{n} \sup_{I \in \mathfrak{J}} |P_n(I) - N_{0,1}(I)| = \frac{2}{\sqrt{2\pi}},$$

which is strictly smaller than $2c_{\infty, \text{BE}}$. Let us note that $c_{\infty, \text{BE}}(\mathfrak{J})$ happens to be twice the modified asymptotic Berry-Esseen constant due to Rogozin [3], and a consideration of the symmetric Bernoulli case shows that a Rogozin type modification of $c_{\infty, \text{BE}}(\mathfrak{J})$ yields the same value $\sqrt{2/\pi}$ as above.

2. RESULTS

For $P \in \mathcal{P}_3$, we shall use the following asymptotic expansion due to Esseen [1]:

$$(3) \quad F_n(x) = \Phi(x) + \frac{h}{\sigma\sqrt{n}}\psi_n(x)\varphi(x) - \frac{\alpha}{6\sigma^3\sqrt{n}}\varphi''(x) + o\left(\frac{1}{\sqrt{n}}\right) \quad (n \rightarrow \infty)$$

uniformly in $x \in \mathbb{R}$, where, if $h > 0$, ψ_n is a certain $h/(\sigma\sqrt{n})$ -periodic function: If $a \in \mathbb{R}$ is such that $P(a + h\mathbb{Z}) = 1$, then $\psi_n(x) = 1/2 - \text{frac}(x\sigma\sqrt{n}/h - an/h)$, where $\text{frac}(t) := t - [t]$ for $t \in \mathbb{R}$. As observed in [2], it then follows that

$$\lim_{n \rightarrow \infty} \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| = \frac{1}{\sqrt{2\pi}} \left(\frac{h}{2\sigma} + \frac{|\alpha|}{6\sigma^3} \right).$$

The analogous result for interval probabilities is:

Theorem 2.1. *Let $P \in \mathcal{P}_3$. Then*

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt{n} \sup_{I \in \mathfrak{J}} |P_n(I) - N_{0,1}(I)| \\ = \frac{1}{\sqrt{2\pi}} \begin{cases} \frac{h}{\sigma} & \text{if } |\alpha| \leq h\sigma^2, \\ \frac{h}{2\sigma} + \frac{|\alpha|}{6\sigma^3} + \frac{|\alpha|}{3\sigma^3} \exp\left(-\frac{3}{2}\left(1 - \frac{h\sigma^2}{|\alpha|}\right)\right) & \text{if } |\alpha| > h\sigma^2. \end{cases} \end{aligned}$$

Proof. We may assume $\alpha \geq 0$, $\sigma = 1$, and $\mu = 0$. On the right of (1), we may omit the absolute value signs, since for a function f satisfying $f(x, y) = -f(y, x)$, we have $\sup_{x, y} |f(x, y)| = \sup_{x, y} f(x, y) \vee (-f(x, y)) = \sup_{x, y} f(x, y) \vee \sup_{x, y} f(y, x) = \sup_{x, y} f(x, y)$. Thus, using (3), we get (reading backwards to prove existence of the limits)

$$\begin{aligned} \lim_{n \rightarrow \infty} 2\sqrt{n} \sup_{I \in \mathfrak{J}} |P_n(I) - N_{0,1}(I)| \\ = \lim_{n \rightarrow \infty} \sup_{x, y \in \mathbb{R}} \left(2h(\psi_n(x)\varphi(x) - \psi_n(y)\varphi(y)) - \frac{\alpha}{3}(\varphi''(x) - \varphi''(y)) \right) \\ = \sup_{x, y \in \mathbb{R}} \left(h\varphi(x) - \frac{\alpha}{3}\varphi''(x) + h\varphi(y) + \frac{\alpha}{3}\varphi''(y) \right) \\ = \sup_{y \in \mathbb{R}} f(y) \end{aligned}$$

where $f(y) := \frac{1}{\sqrt{2\pi}}(h + \alpha/3) + h\varphi(y) + \alpha\varphi''(y)/3$ and where, if $h > 0$, the second equality above is true since $\sup_{x \in \mathbb{R}} \psi_n(x) = 1/2$, $\inf_{x \in \mathbb{R}} \psi_n(x) = -1/2$, the period length of ψ_n tends to zero, and the function occurring on the right is continuous.

If $\alpha = 0$, then f is maximized at $y = 0$ with the value $f(0) = 2h/\sqrt{2\pi}$. Let now $\alpha > 0$. Then $f'(y) = h\varphi'(y) + \alpha\varphi'''(y)/3 = -\alpha y\varphi(y)(y^2 - 3 + 3h/\alpha)/3$. If also $\alpha \leq h$, then $\{f' > 0\} =]-\infty, 0[$, and f is again maximized at $y = 0$, with the same value as above. If instead $\alpha > h$, then $\{f' > 0\} =]-\infty, -y_0[\cup]0, y_0[$ with $y_0 := \sqrt{3 - 3h/\alpha}$, and the maximal value of f is

$$f(\pm y_0) = \frac{1}{\sqrt{2\pi}} \left(h + \frac{\alpha}{3} + \frac{2\alpha}{3} \exp\left(-\frac{3}{2}\left(1 - \frac{h}{\alpha}\right)\right) \right). \quad \square$$

For completeness we present a short proof of a classical fact needed below:

Lemma 2.2 (von Mises [5]). *Let $P \in \text{Prob}(\mathbb{R})$ be discrete and let $\eta \in [0, \infty[$ be such that $0 < |x - y| < \eta$ implies $P(\{x\})P(\{y\}) = 0$. Let $s \in [1, \infty[$. Then*

$$(4) \quad \eta\beta_s \leq 2\beta_{s+1}.$$

Finite equality $\eta\beta_s = 2\beta_{s+1} < \infty$ holds iff $s = 1$ and $P = \lambda\delta_x + (1 - \lambda)\delta_{x+\eta}$ for some $\lambda \in [0, 1]$ and $x \in \mathbb{R}$, or $s > 1$ and $P \in \bigcup_{x \in \mathbb{R}} \{\delta_x, (\delta_x + \delta_{x+\eta})/2\}$.

Proof. We may assume $\eta > 0$ and $P \in \mathcal{P}_{s+1} \setminus \{(\delta_x + \delta_y)/2 : x, y \in \mathbb{R}\}$, all other cases being trivial, and then also $\mu = 0$. Let X, Y be independent with law P . Then $|X - \mu| = |X|$ is not constant almost surely, and hence the map $[1, s+1] \ni t \mapsto \log \beta_t \in \mathbb{R}$ is strictly convex by Hölder's inequality, yielding $\beta_{s+1}/\beta_s \geq \beta_2/\beta_1$, with equality iff $s = 1$. We also have $2\beta_2 = \mathbb{E}(X - Y)^2 \geq \eta\mathbb{E}|X - Y| \geq \eta\mathbb{E}|X| = \eta\beta_1$, with equality in the first inequality iff P is of the form $\lambda\delta_x + (1 - \lambda)\delta_{x+\eta}$, in which case equality holds throughout. Hence (4) holds, with the discussion of equality. \square

Corollary 2.3. *Identity (2) is true, with the $\sup_{P \in \mathcal{P}_3}$ attained precisely for $P \in \{(\delta_x + \delta_y)/2 : x, y \in \mathbb{R}, x \neq y\}$.*

Proof. Let $P \in \mathcal{P}_3$ and let L denote the limit in Theorem 2.1. If $|\alpha| \leq h\sigma^2$, then Lemma 2.2 with $s = 2$ and $\eta = h$ yields $\sqrt{2\pi}L = h/\sigma \leq 2\beta_3/\sigma^3$, with equality in the last step iff P is of the form $(\delta_x + \delta_y)/2$. If $|\alpha| > h\sigma^2$, then $\sqrt{2\pi}L < h/(2\sigma) + |\alpha|/(6\sigma^3) + |\alpha|/(3\sigma^3) < |\alpha|/\sigma^3 \leq \beta_3/\sigma^3$. \square

REFERENCES

- [1] Esseen, C.-G. (1945). Fourier analysis of distribution functions. *Acta Math.* **77**, 1–125.
- [2] Esseen, C.-G. (1956). A moment inequality with an application to the central limit theorem. *Skand. Aktuarietidskr.* **39**, 160–170.
- [3] Rogozin, B.A. (1960). A remark on Esseen's paper: "A moment inequality with an application to the central limit theorem". *Theory Probab. Appl.* **5**, 114–117.
- [4] Shevtsova, I.G. (2011). On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands. *Preprint*, arXiv:1111.6554v1
- [5] von Mises, R. (1938). An inequality for the moments of a discontinuous distribution. *Skand. Aktuarietidskr.* **22**, 32–36.

UNIVERSITÄT TRIER, FACHBEREICH IV – MATHEMATIK, 54296 TRIER, GERMANY
E-mail address: {dinev, mattner}@uni-trier.de